CATÓLICA
UNIVERSIDADE CATÓLICA PORTUGUESA | PORTO
Faculdade de Economia e Gestão

DOCUMENTOS DE TRABALHO — WORKING PAPERS

ECONOMIA — ECONOMICS

Nº 14/2007

DISCRIMINATING MEAN AND VARIANCE SHIFTS

Carlos Santos

Universidade Católica Portuguesa (Porto)

Discriminating mean and variance shifts

Carlos Santos*
Faculdade de Economia e Gestão and CEGE
Universidade Católica Portuguesa

September 25, 2007

Abstract

A two-stage procedure based on impulse saturation is suggested to distinguish mean and variance shifts. The resulting zero-mean innovation test statistic has a non standard distribution, with a nuisance parameter. Hence, simulation-based critical values are provided for some cases of interest. Monte Carlo evidence reveals the test has good power properties to discriminate mean and variance shifts identified through the impulse saturation break test.

JEL Classification Codes: C12; C15; C22; C52
Keywords: breaks; mean shift; variance shift; impulse saturation; nuisance parameter

1 Introduction

Santos, Hendry and Johansen (2007) have established that a general-to-specific strategy is feasible to select from a set of T candidate indicator variables, one for each observation. This principle came to be known as impulse saturation, and is a key result in model selection. Doornik and Sprudsz (2007) and Nielsen and Johansen (2007) have further developed the procedure. Such an initial model cannot be estimated from the outset, so subset selection is used (where the subsets are sample partitions either in halves, thirds, etc), followed by searches across the union of the terminal models. For a split of $T/2$, this entails saturating half the sample and storing the significant indicators, and then examining the other half. Under the null hypothesis that no indicator matters, the impulse saturation procedure is shown to have the correct null rejection frequencies (NRFs) precluding overfitting, independently of the number of splits used for the subsets. For individual tests conducted on each indicator at a significance level α, the average retention rate is αT. The asymptotic distribution of the post-selection estimators of the mean and variance, in a location-scale model

*csantos@porto.ucp.pt

with IID errors is derived in Santos et. al (2007), and extensive Monte Carlo evidence confirms the theoretical results.

Santos and Hendry (2007) and Santos (2006) have further extended this result, by showing that under the alternative of dummies in the DGP the procedure has power against both mean and variance shifts. This is true both in the canonical i.i.d. case as well as in certain classes of dynamic models (Santos and Hendry, 2006).

In this paper we deepen the research on impulse saturation break tests by developing a procedure to discriminate between mean and variance shifts, since impulse saturation break tests have power against both. This new procedure is of the utmost importance for practitioners, since the impulse saturation break test alone can only reveal the existence of a shift, providing no insight into its nature.

The paper is organized as follows. The next section discusses the zero-mean innovation test and the presence of a nuisance parameter on the null distribution. Section 3 presents simulation-based critical values for some cases of interest. Section 4 discusses Monte Carlo evidence on power. Section 5 concludes.[1]

2 Discriminating breaks in the mean and in the variance in location-scale models

Having concluded from Santos and Hendry (2007) that impulse saturation tests have power both against mean shifts and variance shifts, it is of interest to provide the practitioner with a tool to help her distinguish between these. Let S_{α_1} be the set of indicators' coefficients estimates retained from impulse saturation, using some partition of the sample (say, $T/2$). Let $\hat{\psi}_t^*$ be a typical element of S_α. Suppose that we wish to study the distribution of the test statistic for the null hypothesis:

$$\mathsf{H}_0 : \tau = 0 \tag{1}$$

versus the alternative:

$$\mathsf{H}_1 : \tau \neq 0 \tag{2}$$

where $\tau = \mathsf{E}\left[\hat{\psi}^*\right]$. The relevant test statistic is the method of moments estimator:

$$\bar{\hat{\psi}}^* = \frac{1}{G(\alpha_1; T)} \sum_{t \in S_{\alpha 1}} \hat{\psi}_t \tag{3}$$

Then, $\sum_{t \in S_{\alpha_1}} \hat{\psi}_t$ is the sum of the estimated coefficients of the retained indicators from the impulse saturated model. $G(\alpha_1; T) = \#S_{\alpha_1}$ that is, the number of elements in the set of significant impulse indicators in the two auxiliary regressions. For the reasons discussed above, this number depends on α_1 and on T.

[1] All codes were written in Ox 3.4 (Doornik, 2001) and are available from the author on request.

For each α_1 and T, we can write more simply:

$$\bar{\hat{\psi}}^* = \frac{1}{G} \sum_{t \in S_{\alpha_1}} \hat{\psi}_t \tag{4}$$

In short, $\bar{\hat{\psi}}^*$ is the sample mean of the estimated coefficients associated with statistically significant impulse indicator variables.

Then, the following statistic is used to test H_0 versus H_1:

$$\mathsf{t}_{\alpha_1;T} = \frac{\bar{\hat{\psi}}^*}{\sqrt{\hat{\mathsf{V}}\left[\bar{\hat{\psi}}^*\right]}} \underset{\mathsf{H}_0}{\sim} D\left(Z;\alpha_1,T\right) \tag{5}$$

where $D\left(Z;\alpha_1,T\right)$ is some unknown distribution[2] depending on nuisance parameter α_1. The critical values will depend on the significance level used to test H_0, say α_2. However, the particular feature of $D\left(Z;\alpha_1,T\right)$ is that it also depends on the significance level used to retain dummies. Furthermore, under the null, the number of retained dummies will also depend on T. Z in $D\left(Z;\alpha_1,T\right)$ simply represents the vector of observations of z_t.

Whilst it is possible for a distribution of a test statistic to depend on T, say, via degrees of freedom, as with the usual individual significance test in classical regression models, here we also have to take into account the effect of the significance level used for impulse saturation α_1, when testing, at α_2, if the mean of the relevant impulses coefficients' estimators is zero. In the following subsection we study this issue by means of Monte Carlo simulations.

Intuitively, if $\alpha_1 = 0$, and $\alpha_2 > 0$, then $\alpha_1 T = 0$ and $\#S_{\alpha_1} = 0$. Therefore, $\bar{\hat{\psi}}^*$ is not well defined. On the other hand, if only one indicator is retained, then $\mathsf{t}_{\alpha_1;T}$ is also not well defined, since $\hat{\mathsf{V}}\left[\bar{\hat{\psi}}^*\right] = 0$.

In conclusion, α_1 plays a fundamental role in $D\left(Z;\alpha_1,T\right)$. Hence, the critical values to test (1) are dependent on a nuisance parameter, α_1.

3 Simulation-based critical values for $D\left(Z;\alpha_1,T\right)$

We shall obtain critical values for $D\left(Z;\alpha_1,T\right)$ by simulation methods.[3] In particular, we shall consider the cases of samples of sizes $T = 300$, $T = 200$ and $T = 100$. We shall consider values for α_1 in the set $\{0.01; 0.025; 0.05\}$. It is not interesting to consider higher values of α_1, since this would lead, under the null, to a high retention of irrelevant indicators. For impulse saturation,

[2] We are well aware that the $\hat{\psi}_t$ are not independent. However, the presence of a nuisance parameter will lead us to obtain simulation-based critical values, so this fact is irrelevant for the aim of this paper.

[3] Trying to get nuisance parameter free critical values through bootstrapping could be another approach.

$\boldsymbol{\alpha}_1 \setminus$ Q	0.95	0.975	0.9875	0.995
0.01	4.563	5.07	5.482	5.959
0.025	3.539	4.241	4.808	5.583
0.05	2.269	2.679	3.157	3.636

Table 1: Empirical Quantiles of the observed values of the test statistic (5), $T = 300$

$\alpha_1 = 0.01$ would in fact be recommended, to avoid excess spurious dummies. $\alpha_1 = 0.025$ is also looked at in some settings, where retention of some impulses in the saturation stage is important. However, $\alpha_1 = 0.05$ is only included here for reference, as it would not be advised to use it in a saturation procedure, for relevant sample sizes.

We draw $M = 10000$ samples of size T from $\mathsf{N}[0, 1]$. That is, we assume that $\boldsymbol{\varepsilon}$ is a $(T \times 1)$ vector of typical element ε_t. Further, we consider that $\mu = 0, \forall t$. Hence, we obtain $M = 10000$ samples of size T of the z_t process. For each sample, we estimate two regression models. We test the statistical significance of each indicator in each of two regression models, and retain the relevant ones. This retention implies that their estimated coefficients, for each sample, are stored in column vector of dimensions $(G(\alpha_1; T) \times 1)$. Evidently, the number of rows in the column vector might differ from one replication to another (that is, from one draw to some other), as different random numbers are being generated at each loop, albeit coming from the same distribution, allowing for the possibility that the number of aberrant observations or outliers differs between draws.

For each sample, the mean of the column vector Z is:

$$\bar{\hat{\psi}}^* = \frac{1}{G(\alpha_1; T)} \sum_{t \in S_{\alpha_1}} \hat{\psi}_t \tag{6}$$

We compute for each iteration (that is, for each of the $M = 10000$ samples) the value of the statistic $\mathsf{t}_{\alpha_1;T}$. With 10000 values for the relevant statistic, obtained under the null, we look for the quantiles of interest. Given that the hypothesis test is two-sided, we look for the quantiles $\{0.95; 0.975; 0.9875; 0.995\}$. These match, respectively, significance levels for α_2 of 0.1, 0.05, 0.025 and 0.01. Tables (1), (2) and (3) report the results of the empirical quantiles, over 10000 values of the observed statistic under the null. Table (1) refers to a sample size $T = 300$, table (2) refers to $T = 200$ and table (3) refers to a sample size of $T = 100$.

In each of the three tables, for a given empirical quantile, the critical value differs with the significance level used in the previous stage of impulse saturation and relevant dummies retention. The differences between the tables, for corresponding cells, highlight the importance of adjusting critical values for the sample size as well. Tables (1)-(3) should be read as follows: in case a researcher wishes to conduct a test on the expected value of the estimators of the coefficients of retained indicators, when a significance $\alpha = 0.01$ was used on the

$\boldsymbol{\alpha}_1 \setminus$ Q	0.95	0.975	0.9875	0.995
0.01	4.276	4.777	5.207	5.58
0.025	3.906	4.567	5.069	5.47
0.05	2.583	3.142	3.592	4.125

Table 2: Empirical Quantiles of the observed values of the test statistic (5), $T = 200$

$\boldsymbol{\alpha}_1 \setminus$ Q	0.95	0.975	0.9875	0.995
0.01	3.978	4.38	4.809	5.212
0.025	3.994	4.496	4.852	5.289
0.05	3.24	3.975	4.341	4.698

Table 3: Empirical Quantiles of the observed values of the test statistic (5), $T = 100$

impulse saturation stage, then, if he wishes to keep using a significance level $\alpha = 0.01$ for the two-sided zero-mean innovation test, the critical values are 5.212 for a sample size of $T = 100$, 5.58 for a sample size of $T = 200$, and 5.959 for a sample size of $T = 300$.

4 MC evidence on discriminating mean and variance shifts

In case a shift took place, it is of interest to know whether it was mean or a variance shift. For this purpose we shall use the zero mean innovation test and the simulation-based critical values of the previous subsection.

In particular, consider the following DGPs:

(I) Mean Shift

$$z_t = \begin{Bmatrix} \varepsilon_t \Leftarrow T < T_1 \\ d + \varepsilon_t \Leftarrow T \geq T_1 \end{Bmatrix} \tag{7}$$

(II) Variance Shift

$$z_t = \begin{Bmatrix} \varepsilon_t \Leftarrow T < T_1 \\ \sqrt{\theta}\varepsilon_t \Leftarrow T \geq T_1 \end{Bmatrix} \tag{8}$$

where, in both cases, it is assumed that $\varepsilon_t \sim \mathsf{IN}\,[0, 1]$. Hence, we are assuming that $\sigma^2_{z,0} = 1$. Further, we make the simplification $\mu_{2,0} = 0$, referred to in the previous section. We allow d to take values from $\{1; 2; 2.5; 3; 4\}$, and θ to take values from $\{2; 3; 4; 5\}$. For each sample size, it is of relevance to notice that $r = \frac{T - T_1 + 1}{T} = 0.2$, where r is the fraction of observations in the break period. Breaks are assumed to occur at the end of the sample.

Tables (4)-(6) report the empirical NRFs for the case of a level shift. Each table refers to a sample size of 300, 200 and 100, respectively. For every d, T, and α_2 the significance level used for retention in the saturated model was

$\boldsymbol{\alpha}_2 \setminus d$	1	2	2.5	3
0.01	0.1241	0.8475	0.9395	0.9203
0.025	0.1729	0.9189	0.9881	0.9942
0.05	0.2253	0.9536	0.9968	0.9996

Table 4: Empirical Rejection Frequencies for a mean shift: $T = 300$, DGP in (7)

$\boldsymbol{\alpha}_2 \setminus d$	1	2	2.5	3
0.01	0.1218	0.7963	0.9446	0.9669
0.025	0.166	0.8605	0.975	0.9932
0.05	0.2257	0.9062	0.9917	0.9993

Table 5: Empirical Rejection Frequencies for a mean shift: $T = 200$, DGP in (7)

$\alpha_1 = 0.01$. This implies a critical value of 5.959 for a sample of size 300, a critical value of 5.58 for $T = 200$, and 5.212 for $T = 100$.

Firstly, for level shifts of magnitude greater or equal to $2\sqrt{\sigma_{zz}}$ (which amounts to say, given that we assumed $\sigma_{zz} = 1$, $d \geq 2$), the power to reject the hypothesis that the mean of the estimators of the coefficients of the retained impulses is zero is high, even when this test is conducted at $\alpha_2 = 0.01$).

So the zero-mean innovation $\mathtt{t}$-test is capable of rejecting the null in the case of a level shift. Notwithstanding, it remains to be checked how the test performs with retained indicators due only to a variance shift.

Tables (7)-(9) report the results of the empirical NRFs for the case of a variance shift. Sample sizes are $T = 300$, $T = 200$ and $T = 100$, respectively. Again, $\alpha_1 = 0.01$ is assumed throughout. The same empirical quantiles used as references for a significance level $\alpha = 0.01$, in tables (7)-(9), are used here.

In contrast with tables (4)-(6), rejection frequencies of the null are quite low. Furthermore, for $d \geq 2$, rejection frequencies in tables (7)-(9) are well below those in (4)-(6). Most importantly, if one wishes a clear distinction between level and variance shifts, the use of $\boldsymbol{\alpha}_2 = 0.01$ for the zero-mean innovation $\mathtt{t}$-test is advised.

$\boldsymbol{\alpha}_2 \setminus d$	1	2	2.5	3
0.01	0.0843	0.6396	0.8676	0.9476
0.025	0.1266	0.7159	0.9122	0.9777
0.05	0.1852	0.7788	0.946	0.9921

Table 6: Empirical Rejection Frequencies for a mean shift: $T = 100$, DGP in (7)

$\boldsymbol{\alpha}_2 \setminus \theta$	2	3	4	5
0.01	0.038	0.0342	0.0247	0.0175
0.025	0.0617	0.055	0.0394	0.0286
0.05	0.0875	0.0773	0.0574	0.0429

Table 7: Empirical Rejection Frequencies for a variance shift: $T = 300$, DGP in (8)

$\boldsymbol{\alpha}_2 \setminus \theta$	2	3	4	5
0.01	0.0445	0.0717	0.0648	0.0528
0.025	0.0712	0.0935	0.0848	0.0706
0.05	0.1124	0.1298	0.1173	0.098

Table 8: Empirical Rejection Frequencies for a variance shift: $T = 200$, DGP in (8)

5 Conclusion

A test to distinguish mean and variance shifts was presented and Monte Carlo evidence on its empirical power provided. Such evidence suggests the zero-mean innovation test has good power do discriminate between mean and variance shifts.

$\boldsymbol{\alpha}_2 \setminus \theta$	2	3	4	5
0.01	0.0422	0.0911	0.1185	0.1219
0.025	0.0716	0.132	0.1549	0.1559
0.05	0.117	0.1843	0.1998	0.1975

Table 9: Empirical Rejection Frequencies for a variance shift: $T = 100$, DGP in (8)

References

Doornik, J. A. (2001) OX, An Object Oriented Matrix Programming Language, London, Timberlake Consultants

Doornik, J. A. and U. Sprudsz (2007) Econometric Modelling when there are more Variables than Observations, Nuffield College, mimeo

Nielsen, B. and S. Johsanen (2007) Saturation by Indicators in Autoregressive Models, paper presented at the Conference in Honour of David F. Hendry, Oxford (http://www.economics.ox.ac.uk/hendryconference/Papers/Nielsen_DFHVol.pdf)

Santos, C. (2006) Structural Breaks and Outliers in Economic Time Series: Modeling and Inference, D.Phil. Thesis, Oxford University

Santos C. and D. F. Hendry (2006) Saturating Autoregressive Models, Notas Económics, 24 (invited contribution)

Santos C. and D. F. Hendry (2007) Impuse Saturation Break Tests, Economic Letters, *Forthcoming*

Santos, C., Hendry, D. F. and S. Johansen (2007) Automatic Selection of Indicators in a Fully Saturated Regression, Computational Statistics, *Forthcoming*